

Online Appendices to “Education and Geographical Mobility: The Role of the Job Surplus”

Michael Amior

Contents

A CPS sample description and supplementary estimates	2
A.1 Sample description	2
A.2 Historical changes in mobility differentials	2
A.3 Breakdown of migration by reported reasons for moving	2
A.4 Robustness to top earner restriction	6
B Contribution of returning students	7
C Predictive power of PSID subjective mobility costs	8
D Subjective mobility cost estimates from SCE	10
E Robustness of net migration patterns	11
F Wage returns of non-movers	13
G Linear regressions on abstract task content	14
H Theoretical proofs and derivations	15
H.1 Derivation of equation (5) in Section 3.2	15
H.2 Proof of Proposition 1 in Section 3.4	16
H.3 Proof of Proposition 2 in Section 3.5	17
H.4 Expression for annuitized cost $m(\mu \varepsilon)$ in Section 4.1	18
H.5 Proof of Proposition 3 in Section 4.3	19
H.6 Proof of Proposition 4 in Section 4.3	20
H.7 Derivation of equation (15) in Section 6	21
H.8 Proof of Proposition 5 in Section 7	22
I Additional details on theoretical extensions	24
I.1 Endogenous search intensity	24
I.2 Home bias in utility	25
I.3 Idiosyncratic amenity draws	26
I.4 Common amenity effects	27

A CPS sample description and supplementary estimates

A.1 Sample description

The Current Population Survey (CPS) estimates in the main text are based on March waves between 1999 and 2018, taken from the IPUMS database (Flood et al., 2018). The March waves include the Annual Social and Economic Supplement, which reports whether respondents lived in a different state 12 months previously. Since 1999, individuals have also reported their primary reason for moving.

I consider five education groups: high school dropouts (less than 12 years of schooling), high school graduates (12 years), some college (less than an undergraduate degree), and undergraduate and postgraduate degree-holders. Potential labor market experience is defined as age minus years of education minus 6 (or age minus 16, whichever is smaller). I set years of schooling to 13 for individuals with some college but no degree, 14 for associate degrees, 16 for undergraduate, 18 for Master's, 19 for professional and 21 for doctorate degrees. I restrict the sample to individuals with 2-30 years of experience at the survey date: this excludes people with less than one year of experience at the time of moving. I also restrict attention to individuals living in the US one year previously.

Kaplan and Schulhofer-Wohl (2012) show there are inconsistencies in the CPS's procedure for imputing migration status in non-response cases: the imputed data artificially inflate the cross-state migration rate between 1999 and 2005. As it happens, the non-response rate for migration status varies little with education: 13% for college graduates and 14% for non-graduates. I choose to drop all these observations.

A.2 Historical changes in mobility differentials

The CPS analysis in the main text is restricted to the period 1999-2018, for which I have information on reasons for moving. But the mobility gap between education groups goes back many decades. In Panel A of Figure A1, I plot annual cross-state migration rates using CPS March waves from 1964 to 2018.

As is well known, migration rates have declined over this period: see e.g. Molloy, Smith and Wozniak (2011); Kaplan and Schulhofer-Wohl (2017). But this decline was fairly uniform across education groups. The ratio of graduate to non-graduate mobility has mostly hovered around 1.4 (Panel B), and there is no clear upward or downward trend over the period as a whole.

A.3 Breakdown of migration by reported reasons for moving

In Table A1, I present detailed disaggregations of cross-state and cross-county migration in the CPS by reported reason for moving. The first column gives the percentage of the full sample

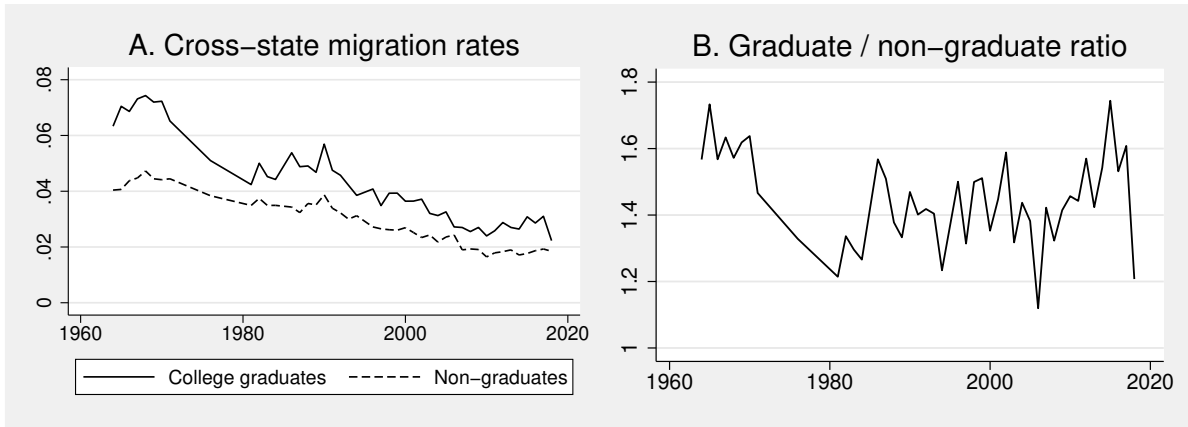


Figure A1: Annual cross-state migration rates by education (1964-2018)

Panel A reports annual rates of cross-state migration over time among individuals with 2-30 years of potential experience in the CPS, separately for college graduates and non-graduates. Panel B shows the ratio of the two. I exclude all individuals living abroad one year previously, and I also exclude observations for which the CPS has imputed migration status: see Appendix [A.1](#) for further details. No migration information is available in 1972-5, 1977-80, 1985 and 1995.

who changed state (in the previous 12 months) for each recorded reason, and the second column expresses these numbers as a percentage of cross-state migrants. The final two columns repeat this exercise for cross-county moves: these consist of both moves across states and across counties within states.

The bottom row shows that, each year, 2.4% of the sample move across states and 5.4% across counties. About half of cross-state moves are motivated by a specific job, compared with a third of cross-county moves. These are mostly due to a job change or transfer, but some workers also report commuting reasons. The commuting motivation can be interpreted in the context of a long-distance match: after accepting a distant job offer (with a long associated commute), the worker eventually changes residence. In contrast, it is rare to move to *look* for work without a job lined up. This sort of speculative job search accounts for just 5% of cross-state and 4% of within-state moves. This is unsurprising: moving without a job in hand is a costly and risky strategy. In terms of non-job migration, family and housing motivations account for most moves.

In Table [A2](#), I report the cross-state and cross-county migration rates (in columns 1 and 3 of Table [A1](#)) separately by education group: high school dropouts (HSD), high school graduates (HSG), some college (SC), undergraduate degree (UG) and postgraduate (PG). As before, the first row reports the rate of job-motivated migration. Notice the (positive) education slope is steeper in proportional terms for cross-state than cross-county moves. I offer a rationale for this result in Section [3.7](#): to the extent that cross-state migration is more costly, education differences in match quality returns and job surplus should matter more. Also, as in Figure [2](#) in the main text, better educated individuals make fewer speculative moves to “look for work”.

On aggregate, there is a mild negative education gradient in non-job migration, which is stronger for cross-county moves. This effect is driven by a broad range of motivations: mostly to “establish own household”, “other family reasons”, “cheaper housing”, “other housing rea-

Table A1: Breakdown of primary reasons for moving

Primary reason	State moves		County moves	
	% full sample	% state migrants	% full sample	% county migrants
DUE TO SPECIFIC JOB	1.11	47.02	1.90	35.53
New job or job transfer	0.94	39.82	1.38	25.82
Easier commute	0.05	2.30	0.33	6.23
Other job reasons	0.12	4.91	0.19	3.48
LOOK FOR WORK	0.13	5.48	0.22	4.03
NON-JOB REASONS	1.12	47.49	3.23	60.44
Family	0.55	23.48	1.40	26.12
Change in marital status	0.10	4.42	0.33	6.08
Establish own household	0.09	3.79	0.36	6.76
Other family reasons	0.36	15.28	0.71	13.28
Housing	0.26	10.86	1.16	21.75
Want to own home	0.04	1.69	0.27	5.04
New or better housing	0.06	2.56	0.35	6.47
Cheaper housing	0.06	2.56	0.22	4.19
Other housing reasons	0.10	4.05	0.32	6.04
Environment	0.13	5.01	0.27	4.91
Better neighborhood	0.04	1.50	0.15	2.71
Climate, health, retirement	0.08	3.51	0.12	2.20
Attend/leave college	0.10	4.39	0.21	3.92
Other reasons	0.09	3.75	0.20	3.74
ALL REASONS	2.35	100	5.35	100

This table presents migration rates by primary reason in CPS March waves between 1999 and 2018. The first column reports the percentage of the *full sample* who changed state, for each given reason, over the previous twelve months. The second column expresses these numbers as a percentage of *state-movers*. The final two columns repeat the exercise for cross-county moves. I include individuals moving because of foreclosure or eviction in the CPS's "other housing reasons" category; and I include individuals moving because of natural disasters in the "other reasons" category. See Appendix [A.1](#) for sample details.

Table A2: Education gradients by primary reasons for moving

Primary reason	State moves					County moves				
	HSD	HSG	SC	UG	PG	HSD	HSG	SC	UG	PG
<u>DUE TO SPECIFIC JOB</u>	0.43	0.74	0.92	1.63	2.20	0.93	1.37	1.75	2.72	3.03
New job or job transfer	0.33	0.59	0.77	1.42	1.95	0.58	0.91	1.21	2.09	2.46
Easier commute	0.04	0.05	0.05	0.06	0.08	0.22	0.29	0.36	0.42	0.33
Other job reasons	0.06	0.10	0.11	0.15	0.17	0.13	0.17	0.18	0.22	0.23
<u>LOOK FOR WORK</u>	0.21	0.15	0.11	0.12	0.07	0.39	0.25	0.19	0.17	0.10
<u>NON-JOB REASONS</u>	1.15	1.18	1.16	1.11	0.82	3.42	3.45	3.48	3.07	2.16
Family	0.63	0.64	0.58	0.46	0.36	1.57	1.62	1.49	1.19	0.82
Change in marital status	0.08	0.12	0.12	0.09	0.09	0.22	0.35	0.37	0.32	0.25
Establish own household	0.11	0.10	0.10	0.07	0.05	0.37	0.42	0.40	0.33	0.19
Other family reasons	0.44	0.42	0.37	0.29	0.22	0.97	0.85	0.72	0.54	0.38
Housing	0.29	0.27	0.23	0.27	0.23	1.27	1.22	1.16	1.17	0.92
Want to own home	0.03	0.04	0.04	0.05	0.05	0.20	0.25	0.27	0.34	0.27
New or better housing	0.08	0.07	0.05	0.06	0.04	0.36	0.36	0.34	0.38	0.28
Cheaper housing	0.08	0.08	0.05	0.05	0.04	0.31	0.27	0.22	0.17	0.13
Other housing reasons	0.10	0.09	0.09	0.10	0.10	0.40	0.35	0.33	0.28	0.24
Environment	0.16	0.15	0.14	0.11	0.08	0.35	0.30	0.29	0.22	0.15
Better neighborhood	0.06	0.04	0.04	0.02	0.02	0.21	0.16	0.15	0.12	0.08
Climate, health, retirement	0.09	0.09	0.09	0.08	0.05	0.14	0.14	0.12	0.09	0.07
Attend/leave college	0.00	0.05	0.13	0.19	0.10	0.02	0.10	0.34	0.30	0.16
Other reasons	0.08	0.09	0.09	0.10	0.07	0.21	0.22	0.21	0.19	0.12
<u>ALL REASONS</u>	1.79	2.07	2.2	2.86	3.09	4.73	5.07	5.42	5.96	5.29

This table reports cross-state and cross-county migration rates by primary reason for moving (as in Table [A1](#)), but now disaggregated by education. I consider five education groups: high school dropouts (HSD), high school graduates (HSG), some college (SC), undergraduate degree (UG) and postgraduate (PG).

Table A3: Migration rates (%) for all individuals and household top earners

	Specific job		Look for work		Non-job	
	All indiv (1)	Top earners (2)	All indiv (3)	Top earners (4)	All indiv (5)	Top earners (6)
HS dropout	0.43	0.46	0.21	0.21	1.15	1.11
HS graduate	0.74	0.76	0.15	0.15	1.18	1.13
Some college	0.92	0.95	0.11	0.09	1.16	1.18
Undergraduate	1.63	1.67	0.12	0.11	1.11	1.12
Postgraduate	2.20	2.36	0.07	0.06	0.82	0.79

This table reports annual cross-state job migration rates by reported reason for moving, separately for all individuals (identical to Figure 1 and Table A1) and for household top earners, and based on CPS March waves between 1999 and 2018. See Appendix A.1 for further details on sample.

sons” and “better neighborhood”. There are just two non-job motivations with (largely) positive education slopes: the desire to purchase a home and attending or leaving college.

The final row reports total migration rates - for all motivations combined. Notice this is much flatter for cross-county migration. Mechanically, this reflects the flatter (positive) gradient of job-motivated migration, the steeper (negative) gradient of non-job migration, and the greater dominance of non-job motivations for cross-county movers.

A.4 Robustness to top earner restriction

The CPS question on reasons for moving is addressed to *individuals* within households. But of course, migration decisions are made in the context of the household. This ambiguity may yield some problems for interpretation: for example, household dependents may choose to simply report the motivations of the breadwinners. This is most clearly illustrated for children (though they are excluded from my sample): in households with at least one adult moving for a specific job, 80% of under-16s also report moving for the same reason.

To address this concern, I recompute migration rates by reason for moving (and by education), but this time restricting the sample to those individuals with the greatest annual earnings in each household. In households with joint top-earners, I divide the person weights by the number of top-earners. This restriction excludes 44% of the original sample. But as Table A3 shows, it makes little difference to the education slopes of job-specific, speculative or non-job migration. Columns 1, 3 and 5 replicate the cross-state migration rates from Table A2, and these look very similar to the remaining columns which impose the top-earner restriction.

Table A4: Contribution of students to mobility differentials

	HS dropout (1)	HS graduate (2)	Some college (3)	Undergraduate (4)	Postgraduate (5)
% recent students by completed education	39.82	28.33	31.60	16.80	8.37
Migration rate (%)					
Full sample	3.10	4.31	5.23	8.78	9.58
Excluding recent students	2.86	4.09	6.03	8.96	9.71
Observations	4,396	9,977	6,125	3,172	1,244

This table reports annual cross-state migration rates by education group, based on all (annual) PSID waves between 1990 and 1997. Migration rates are constructed using reported state of residence 12 months previously. The first row gives the fraction of the sample who were recently students (in the current annual wave, or in the previous three years). The second row reports cross-state migration rates for the full sample, and the third row reports these rates excluding recent students. The sample consists of all individuals with 2-10 years of potential labor market experience at the end of each 12-month interval.

B Contribution of returning students

In this section, I check whether returning students may be contributing to education differentials in mobility. Table A2 shows that workers who report moving primarily to leave or attend college account for a negligible part of these differentials. But even if this is not the primary stated motivation, it may be an underlying factor for those who report job-related reasons - at least for the young. Indeed, Kennan and Walker (2011) emphasize that a large fraction of long-distance movers in the US are returning to former places of residence; and Kennan (2015) studies the tendency of individuals to return home after studying in another state.

I assess the contribution of returning students using the Panel Study of Income Dynamics (PSID), in annual waves between 1990 and 1997. Similarly to the CPS analysis, I define a migrant as somebody living in a different state 12 months previously. As before, I restrict attention to individuals with at least 2 years of potential experience at the end of each 12-month interval; and for this exercise, I also exclude individuals with more than 10 years. I end my sample in 1997, because the PSID became biennial after than year (this prevents me from tracking migration at annual frequencies).

I report my estimates in Table A4. The first row reports the fraction of individuals in each education group who were recently students (in the current annual wave, or in the previous three). This is largest for high school dropouts (40%) and smallest for individuals with undergraduate and postgraduate degrees (17% and 8% respectively).

The remaining rows report annual cross-state migration rates by education. The second row computes these for the full sample, illustrating the familiar positive education gradient. In the third row, I exclude recent students. This makes little difference to the results; and if

anything, graduate migration rates are now slightly larger. This suggests that recent students are not responsible for the mobility gap in this sample.

Of course, excluding *recent* students does not address the concerns entirely, because ex-students may yet return to their home state several years after completing their education. But as I show in Table 3, excluding return moves more generally (i.e. back to individuals' state of birth) only reinforces the elasticity of cross-state matching to education. To summarize then, the evidence shows that returning students (and return migration in general) cannot account for the mobility gap.

C Predictive power of PSID subjective mobility costs

Since the imputed costs in Section 2.3 are based on the subjective judgments of respondents, there may be doubts over their accuracy. But reassuringly, I show here that these cost measures do have significant predictive power for future migration decisions.

Suppose the instantaneous cross-state matching rate for some individual i is constant within the time interval $t - 1$ to t , and denote this rate as ρ_{Cit} (where the subscript C denotes cross-state matching). The probability of moving within this interval is then:

$$\Pr(\text{Move}_{it} = 1) = 1 - \exp(-\rho_{Cit}) \quad (\text{A1})$$

This motivates a complementary log-log model:

$$\Pr(\text{Move}_{it} = 1) = 1 - \exp(-\exp(\beta_m m_{it-1} + \beta'_X X_{it} + \beta_t)) \quad (\text{A2})$$

where I express ρ_{Cit} as a function of the initial mobility cost m_{it-1} , observable demographics¹ X_{it} , and a full set of year effects β_t . The advantage of this specification is that β_m can intuitively be interpreted as the elasticity of the instantaneous migration rate ρ_{Cit} with respect to m_{it-1} . And assuming a constant hazard, this interpretation is independent of the time horizon associated with the migration variable.

I report my estimates in Table A5. Recall from Section 2.3 that I only observe costs for those who report being “willing” to move for work (see Figure 3). Hence, I begin in columns 1-2 by estimating the elasticity of cross-state migration (within 12-month intervals) to a binary indicator for “willingness” to move for work (measured at the beginning of the interval); and conditional on being willing to move, I then report the elasticity to the imputed mobility cost m_{it-1} in columns 3-5. My sample consists of employed household heads between 1969 and 1972, so I restrict attention to the mobility decisions of these individuals in the annual intervals between 1969 and 1973. For this exercise, I use the log specification for imputed costs (i.e. Panel B of Figure 4): this is the difference between (i) the log of a worker’s stated reservation

¹Specifically: gender, experience and experience squared, and four education indicators (high school graduate, some college, undergraduate and postgraduate).

Table A5: Elasticities of cross-state migration to imputed costs

	Unconditional sample		Conditional sample (willing to move)		
	(1)	(2)	(3)	(4)	(5)
Initial willingness to move	0.915 (0.169)	0.810 (0.191)			
Initial willingness to move \times Graduate		0.360 (0.393)			
Initial imputed cost			-0.288 (0.327)	-1.108 (0.381)	-1.017 (0.420)
Initial imputed cost \times Graduate					-0.256 (0.941)
Initial log wage				-1.051 (0.312)	-1.254 (0.346)
Initial log wage \times Graduate					1.083 (0.721)
Demographic controls, year effects	Yes	Yes	Yes	Yes	Yes
Observations	8,835	8,835	3,811	3,811	3,811
Cross-state migration rate	0.034	0.034	0.047	0.047	0.047

This table reports elasticities of cross-state migration (over 12-month intervals) to subjective costs and wages (at the beginning of each interval), based on the complementary log-log model of equation (A2). I study the response to both a binary indicator of "willingness" to move for work; and conditional on being willing to move, the response to the imputed mobility cost. Imputed costs are measured using the log specification, as in Panel B of Figure 4. In columns 2 and 5, I also allow for interactions between the cost measures and a college graduate dummy. Coefficients should be interpreted as the log point effect of each measure on the instantaneous cross-state migration rate. I can only impute costs for employed household heads between 1969 and 1972, so I restrict attention to the mobility decisions of these individuals in the annual intervals between 1969 and 1973. I exclude those with less than 2 or more than 30 years of potential experience at the end of each interval. Household heads in the PSID are always male, unless there is no male partner present. All specifications control for a full set of year effects and demographic controls, specifically gender, experience and experience squared, and four education indicators (high school graduate, some college, undergraduate and postgraduate). Errors are clustered by individual, and robust standard errors are in parentheses.

for accepting a long-distance offer and (ii) the log of the worker's actual wage.

Column 1 shows that initial "willingness" to move approximately doubles an individual's subsequent cross-state migration rate. An interaction with a college graduate dummy reveals no significant difference in the response by education (column 2).

Conditional on willingness to move, column 3 estimates an elasticity of cross-state migration to imputed costs of -0.3, but it is statistically insignificant. This estimate is presumably attenuated by classical measurement error, but there is also a more systematic problem. If the moving reservation is noisier than a worker's wage, the imputed cost (the difference between the two) will be artificially negatively correlated with wages. But as the model in Section 3 shows, to the extent that wages reflect match quality, workers with higher initial wages are less likely to move. This will bias the estimated effect of imputed costs towards zero. To address

this problem, in column 4, I control additionally for the initial log wage. Both the imputed cost and the wage now take strong negative effects (just as theory predicts), with elasticities of -1.1 and standard errors of 0.3 to 0.4. In column 5, I allow for education heterogeneity in these effects, but the interactions are not statistically significant.

The key message here is that the subjective costs *do* have predictive power for future mobility - which suggests they are informative about the true costs of moving. This reinforces the validity of the evidence in Section 2.3 that subjective mobility costs vary little with education.

D Subjective mobility cost estimates from SCE

In Section 2.3, I impute subjective mobility costs using a set of questions from the PSID in the early 1970s. In this appendix, I impute subjective costs using similar questions from the Job Search Supplement of the Survey of Consumer Expectations (SCE): the cost patterns are very similar. This supplement, developed by Faberman et al. (2022), was administered to a subset of respondents in the October waves of the 2013-9 SCE surveys. The SCE is restricted to household heads, similar to my PSID sample. See the online supplement to Faberman et al. (2022) for further details.

I rely on two questions. First, respondents were asked for their basic reservation wage:

- (1) “Suppose someone offered you a job today. What is the lowest wage or salary you would accept (before taxes and deductions) for the type of work you are looking for?” [I denote the answer as w_R , expressed in hourly wages]

Unusually, this question was posed to *all* individuals (both with and without jobs) who reported being open to a new/additional job; only the self-employed were excluded. Second, respondents were asked:

- (2) “Suppose you were offered a job today that paid w_R . Would you accept this job if it required you to relocate to another city or state? ... By what percentage would the wage have to be higher, if at all, for you to relocate?” [Denote the answer as ψ_R , where $\psi_R \geq 0$]

About half of respondents express a willingness to relocate for work (and report a ψ_R) in (2), very similar to the PSID (see Panel A of Figure 3). And also similar to the PSID, this share varies little by education: 52% for high school workers, 56% with some college, 55% for undergraduate degree-holders, and 52% for postgraduates.

Among those who do answer (2), I can impute a dollar reservation for a long-distance offer as $(1 + \psi_R) w_R$; and this allows me to compute cost measures which are comparable to those of Figure 4. Specifically, the dollar cost specification (from Panel A) is the difference between

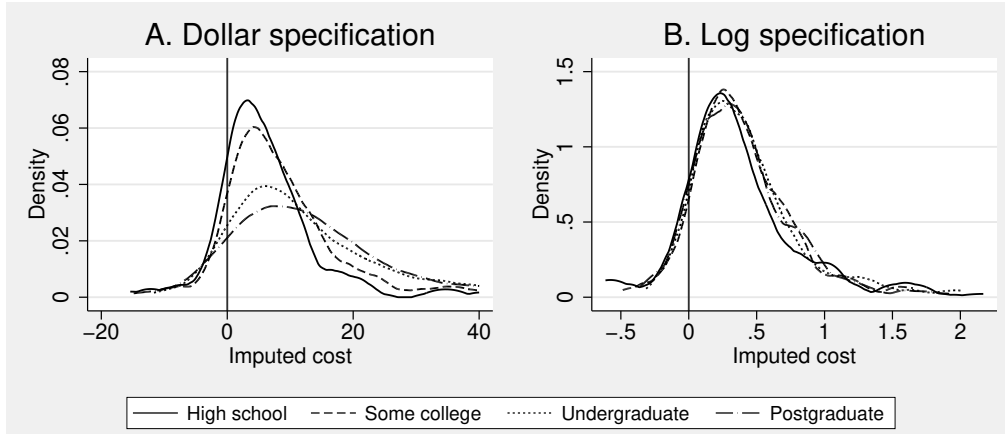


Figure A2: Distribution of imputed costs from SCE

This figure plots kernel distributions of the imputed costs of moving. Conditional on expressing willingness to move, household heads (in the SCE, 2013-19) report their reservation wage for accepting a long-distance offer. In the dollar specification (Panel A), I impute the costs as the difference between this reservation wage and the worker's current wage (in 2015 dollars), where wages are measured in hourly terms. In the log specification (Panel B), I take differences between the log reservation and log current wage. I drop the top and bottom 2% of imputed costs within education groups. For the dollar specification, I only show the support up to \$40 (to make the salient parts of the distribution more visible). I restrict the sample to employed individuals with 1-30 years of experience. The sample consists of 107 individuals with no college, 357 with some college, 609 with undergraduate degrees, and 463 with postgraduate degrees.

the *dollar* long-distance reservation, $(1 + \psi_R)w_R$, and current wage, w . And the log specification (Panel B) is the difference between the *log* long-distance reservation, $\log(1 + \psi_R)w_R$, and current log wage, $\log w$.²

In Figure A2, I plot kernel densities of imputed costs separately by education group. I have grouped all high school workers together, since there are only 6 high school dropouts in the sample (see the figure notes for sample counts). The patterns look very similar to the PSID estimates in Figure 4. For the dollar specification (Panel A), costs are much larger for college graduates: \$19.00 on average (in 2015 dollars), compared to just \$9.48 for non-graduates. But the costs vary little by education in the log specification (Panel B): 0.37 on average for high school workers, 0.39 for some college, 0.42 for undergraduates, and 0.40 for postgraduates. Note that, despite differences in timing and question structure, the log costs are very similar in magnitude to the PSID: the average cost in the SCE is 0.39, compared to 0.37 in the PSID.

E Robustness of net migration patterns

In Table A6, I reproduce the results in Table 2 in the main text, but this time separately for individuals with 2-10 and 11-30 years of potential labor market experience. At least for young college-educated individuals, net mobility is increasing in education (column 2 of Panel A); but crucially, for all experience groups and occupation schemes, the net-gross ratio is still

²An alternative approach would be to use differentials relative to the reservation wage w_R in (1), and not relative to the current wage w . But this approach has an important limitation. For workers who *expect* to relocate for their next job (if their labor market is very thin: i.e. a large σ^e , in the language of my model), their stated reservation w_R in (1) will already account for the cost of moving. In that case, their ψ_R would *understate* this cost.

Table A6: Net cross-state migration rates by education and experience

	Basic			Within 2-digit occupations			Within 3-digit occupations		
	Gross rate %	Net rate %	Net-gross ratio	Gross rate %	Net rate %	Net-gross ratio	Gross rate %	Net rate %	Net-gross ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Panel A: Individuals with 2-10 years of experience</i>									
HS dropout	3.16	0.37	0.12	3.09	1.25	0.41	3.09	1.68	0.54
HS graduate	4.04	0.30	0.08	3.60	0.80	0.22	3.60	1.19	0.33
Some college	4.16	0.32	0.08	3.70	0.79	0.21	3.70	1.24	0.34
Undergraduate	6.42	0.48	0.07	5.79	0.99	0.17	5.79	1.55	0.27
Postgraduate	7.41	0.57	0.08	7.00	1.26	0.18	7.00	1.92	0.27
<i>Panel B: Individuals with 11-30 years of experience</i>									
HS dropout	1.96	0.26	0.13	1.70	0.62	0.37	1.70	0.84	0.49
HS graduate	1.95	0.22	0.11	1.61	0.38	0.23	1.61	0.55	0.34
Some college	2.27	0.22	0.10	1.90	0.43	0.22	1.90	0.65	0.34
Undergraduate	2.31	0.21	0.09	1.98	0.39	0.20	1.98	0.60	0.31
Postgraduate	2.62	0.22	0.08	2.34	0.45	0.19	2.34	0.69	0.29

This table reports annual gross and net cross-state migration rates within education groups, separately for individuals with 2-10 and 11-30 years of potential experience. See notes under Table 2 in the main text for sample details and construction of variables.

Table A7: Net migration rates by education, across metro areas

	Basic			Within 2-digit occupations			Within 3-digit occupations		
	Gross rate %	Net rate %	Net-gross ratio	Gross rate %	Net rate %	Net-gross ratio	Gross rate %	Net rate %	Net-gross ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HS dropout	3.02	0.53	0.17	2.21	1.05	0.47	2.21	1.34	0.60
HS graduate	3.16	0.42	0.13	2.55	0.81	0.32	2.55	1.14	0.45
Some college	3.57	0.44	0.12	3.02	0.91	0.30	3.02	1.32	0.44
Undergraduate	4.03	0.44	0.11	3.64	0.89	0.25	3.64	1.30	0.36
Postgraduate	4.13	0.49	0.12	3.85	0.96	0.25	3.85	1.35	0.35

This table reports annual gross and net migration rates across Metropolitan Statistical Areas (MSAs), within education groups. Variables are constructed in the same way as in Table 2 in the main text (see notes under that table for details), except the sample is now restricted to 2005-17 (MSAs are only identified in these years).

flat or decreasing in education (columns 3, 6 and 9). Thus, mobility differentials between education groups cannot be explained by large net flows to particular states, even within distinct experience categories. This reinforces the general message of Section 2.4 in the main text.

In the main text, I focus on mobility across states. But in Table A7, I reproduce my results for mobility across Metropolitan Statistical Areas (MSAs). Variables are otherwise constructed in the same way as before, except the sample is now restricted to 2005-17 (since MSAs are only identified from 2005 in the ACS data). Again, in every case, the ratio of net to gross migration is decreasing in education (columns 3, 6 and 9), just as in the main text. I conclude that the education differentials in gross flows cannot be explained by large net flows to particular MSAs, both overall and within detailed occupation categories.

Table A8: Net migrant stocks by education

	Basic			Within 2-digit occupations			Within 3-digit occupations		
	Gross share %	Net share %	Net-gross ratio	Gross share %	Net share %	Net-gross ratio	Gross share %	Net share %	Net-gross ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HS dropout	25.69	6.40	0.25	27.51	8.65	0.31	27.51	10.01	0.36
HS graduate	28.85	7.61	0.26	28.71	8.22	0.29	28.71	8.68	0.30
Some college	34.20	8.92	0.26	33.60	9.47	0.28	33.60	9.97	0.30
Undergraduate	43.88	12.24	0.28	43.11	12.56	0.29	43.11	13.09	0.30
Postgraduate	52.08	14.56	0.28	51.67	14.96	0.29	51.67	15.68	0.30

This table reports gross stocks of migrants and net imbalances by education. In this exercise, I define "migrants" as individuals who live outside their state of birth. Column 1 reports the gross share of individuals living outside their birth state. Column 2 reports net imbalances of migrant stocks across states, which I compute in the same way as for net flows in Table 2 in the main text. And column 3 reports the ratio of net imbalances to gross stocks. The remaining columns repeat this exercise within occupation-defined labor markets, just as I do in Table 2. The sample excludes foreign-born individuals, but is otherwise identical to that of Table 2.

One might also be concerned that gross mobility differentials are merely driven by churn: software engineers moving to California, and then returning home. To assess this possibility, I next study *stocks* of migrants, which are not conflated by such churn. In Table A8, I reproduce the analysis above, but now defining "migrants" as people living outside their birth state (rather than recent movers): this excludes return migrants. Looking at column 1, more than half of postgraduate degree-holders live outside their birth state, compared to 26% of high school dropouts. Though local imbalances between worker "imports" and "exports" are increasing in education (column 2), there is little change relative to the gross stock of migrants (column 3). The same is true within occupation-defined markets. This suggests that churn and return migration cannot account for the gross mobility differentials: see also the analysis of return migration in Table 3 in the main text.

F Wage returns of non-movers

Table 4 shows a steep education gradient in the wage returns to same-state job matching. I have argued that these estimates identify education differentials in the match returns parameter, σ^ε . However, one may be concerned that they are driven by mobility across metro areas *within* states. To address this concern, I re-estimate these returns for the sample of non-movers, i.e. individuals who do not change residence at all between $t - 1$ and t . Of course, this restriction introduces its own selection issues, as it excludes individuals who simply moved to a larger house (in the same neighborhood) contemporaneously with the new job match. But as it happens, the sample restriction makes no substantive difference to the results.

I report my estimates in Table A9. Since I do not observe within-state mobility in the 1996 panel, I restrict the sample to observations since 2000. I begin in column 1 by estimating the same-state returns for the post-2000 sample, for individuals who do not change state between $t - 1$ and t (i.e. replicating the exercise of column 1 of Table 4). The estimates are similar to

Table A9: Mean returns to job matching for non-movers

	Mean wage returns		Sample of job transitions	
	Same-state (1)	Non-mover (2)	Same-state (3)	Non-mover (4)
All individuals	0.037 (0.003)	0.035 (0.003)	23,794	21,664
High school	0.023 (0.005)	0.022 (0.005)	8,965	8,145
Some college	0.032 (0.005)	0.029 (0.006)	8,455	7,663
Undergraduate	0.062 (0.009)	0.059 (0.009)	4,559	4,177
Postgraduate	0.073 (0.016)	0.070 (0.017)	1,815	1,679

Table columns do *not* represent regression specifications: reported statistics are mean returns to job matching (for the group specified on the left), accompanied by standard errors (in parentheses). Column 1 reproduces the estimates of same-state returns in column 1 of Table 4 but now restricting the sample to observations since 2000. Column 2 shows mean returns for individuals who do not change residence between $t - 1$ and t . Estimation is otherwise identical to Table 4; see notes under Table 4 for additional details.

those in Table 4

In column 2, I now restrict the sample to individuals who do not change residence between $t - 1$ and t . It turns out this makes little difference. This shows that the education differentials in Table 4 in the main text are not driven by spatial mobility within states. Rather, they appear to genuinely identify differentials in match quality returns, σ^ε . This should not be surprising: as columns 3-4 show, few same-state matches involve changes in residence.

G Linear regressions on abstract task content

In this appendix, I estimate linear regressions which correspond to the line graphs in Figure 5. I estimate the following equation:

$$y_i = \beta_0 + \beta_0^{SC} SC_i + \beta_0^{UG} UG_i + \beta_0^{PG} PG_i + \left(\beta_a + \beta_a^{SC} SC_i + \beta_a^{UG} UG_i + \beta_a^{PG} PG_i \right) \cdot \bar{a}_i + \varepsilon_i \quad (A3)$$

using variation across individual workers i . The outcome y_i is either (i) the mean match returns in worker i 's same-state matches (averaged across the SIPP panel) or (ii) the cross-state share of worker i 's matches. \bar{a}_i is worker i 's mean abstract task content, averaged across the SIPP panel. I interact \bar{a}_i with education effects (dummy variables for some college, undergraduate and postgraduate degree), and also control for these effects separately: the omitted category is high school.

Table A10: Abstract task slopes in match returns and cross-state match shares

	Same-state match returns		Cross-state match share	
	(1)	(2)	(3)	(4)
Abstract content	0.008 (0.002)	0.002 (0.003)	0.003 (0.000)	0.001 (0.000)
Abstract content \times Some college		0.004 (0.004)		0.001 (0.001)
Abstract content \times Undergraduate		0.008 (0.005)		0.003 (0.001)
Abstract content \times Postgraduate		0.019 (0.009)		0.010 (0.002)
Education fixed effects	Yes	Yes	Yes	Yes
Sample	26,541	26,541	76,829	76,829

This table reports OLS estimates of equation (A3). Each observation represents an individual worker. The key regressor of interest is the worker’s mean abstract task content (averaged across the SIPP panel), and this is interacted in columns 2 and 4 with the education effects. All regressions also control for the education effects separately. In columns 1-2, the dependent variable is the mean same-state match return (restricted to job transitions within employment cycles) for each individual, over the SIPP panel. The sample excludes individuals with no qualifying same-state transitions. In columns 3-4, the dependent variable is the cross-state share of each individual’s matches, over the SIPP panel. Robust standard errors are in parentheses.

In Table A10, I report the β_a slope coefficients. Columns 1 and 3 omit the education interactions, and columns 2 and 4 include them. As Figure 5 shows, the β_a slopes are increasing in education for both outcomes. For same-state returns (column 2), the postgraduate slope is significantly steeper (at the 5% level) than the high school slope. For cross-state match share (column 4), both the undergraduate and postgraduate slopes are significantly steeper. To aid interpretation of the β_a coefficients, note that the 10th percentile of \bar{a}_i (across individuals) is 0.7, and the 90th percentile is 6.7.

H Theoretical proofs and derivations

H.1 Derivation of equation (5) in Section 3.2

In this section, I derive the equilibrium distribution G of match quality ε , i.e. equation (5) in the main text. For a worker with initial quality ε , let:

$$\rho(\varepsilon) \equiv \rho_L(\varepsilon) + \rho_C(\varepsilon) \tag{A4}$$

be the overall job matching rate, i.e. the sum of the local and cross-area rates. Let $G(\varepsilon)$ be the share of workers with match quality below ε . Note I define G to account for both employed *and* unemployed workers, with the latter assigned the reservation quality b (which defines the lower bound of the ε distribution). The inflow of workers to this group (i.e. employment below ε or unemployment) must equal the outflow in equilibrium:

$$\delta [1 - G(\varepsilon)] = \rho(\varepsilon) \cdot G(\varepsilon) \quad (\text{A5})$$

The inflow (on the left) consists entirely of workers with initial quality above ε entering unemployment (at separation rate δ), and the outflow (on the right) consists of workers with initial quality in the region $[b, \varepsilon]$ matching with jobs above ε (which occurs at rate $\rho(\varepsilon)$). Rearranging then yields:

$$G(\varepsilon) = \frac{\delta}{\delta + \rho(\varepsilon)} \quad (\text{A6})$$

for $\varepsilon \geq b$, which is equation (5) in the main text.

The corresponding probability density is:

$$g(\varepsilon) = \begin{cases} 0 & \text{if } \varepsilon < b \\ \frac{\delta}{\delta + \rho(b)} & \text{if } \varepsilon = b \\ -\frac{\delta \rho'(\varepsilon)}{[\delta + \rho(\varepsilon)]^2} & \text{if } \varepsilon > b \end{cases} \quad (\text{A7})$$

where the unemployed are treated as receiving match quality b . This is effectively a left-censored distribution, with a discrete probability mass (i.e. the unemployed) at the censored value of b .

H.2 Proof of Proposition 1 in Section 3.4

Proposition 1 states that the odds ratio $\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)}$ is increasing in σ^ε , for given initial match quality ε . Note first that $\rho_L(\varepsilon)$ is invariant to σ^ε : see equation (3). So it is sufficient to show that $\rho_C(\varepsilon)$ is increasing in σ^ε . My argument will revolve around equation (4), which I reproduce here for convenience:

$$\rho_C(\varepsilon) = \pi \lambda \int_\varepsilon^\infty \left[F^\mu \left(\frac{\sigma^\varepsilon}{\sigma^\mu} \int_\varepsilon^{\varepsilon'} \frac{1}{r + \delta + \rho_L(x) + \rho_C(x)} dx \right) \right] dF^\varepsilon(\varepsilon') \quad (\text{A8})$$

Match quality returns σ^ε enters this equation both directly, through the σ^ε term in the numerator, and indirectly, through $\rho_C(x)$ in the denominator. The direct effect is clearly positive for $\rho_C(\varepsilon)$. To prove the proposition then, the main challenge is to deal with the indirect effect. Notice that this indirect effect vanishes as the initial match quality ε becomes very large: this ensures that the $\rho_C(x)$ term in the denominator converges to zero. And therefore, for ε sufficiently large (call this value ε_1), $\rho_C(\varepsilon_1)$ must indeed be increasing in σ^ε .

Next, consider what happens for a slightly smaller value of ε , i.e. $\varepsilon_0 < \varepsilon_1$. The proposition can be demonstrated for ε_0 by contradiction. Suppose that $\rho_C(\varepsilon_0)$ is decreasing in σ^ε at ε_0 . The additional indirect effect (via the denominator) would then be positive for $\rho_C(\varepsilon)$ at ε_0 , so $\rho_C(\varepsilon_0)$ must then be increasing in σ^ε (i.e. we have a contradiction). Consequently, $\rho_C(\varepsilon_0)$ must be increasing in σ^ε at ε_0 .

This same argument can then be iterated, moving down the entire ε distribution. So, $\frac{\partial \rho_C(\varepsilon)}{\partial \log \sigma^\varepsilon}$ must be positive for all ε .

H.3 Proof of Proposition 2 in Section 3.5

Proposition 2 states that: for workers with larger initial match quality ε , the odds ratio $\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)}$ is less sensitive to match returns σ^ε , for sufficiently large mobility cost scale σ^μ . That is, the second derivative $\frac{\partial^2 \left[\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)} \right]}{\partial \log \sigma^\varepsilon \cdot d\varepsilon}$ is negative.

I begin by reproducing the odds ratio from (6) in the main text:

$$\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)} = \pi \frac{\int_{\varepsilon}^{\infty} \left[F^\mu \left(\frac{\sigma^\varepsilon}{\sigma^\mu} \int_{\varepsilon}^{\varepsilon'} \frac{1}{r+\delta+\rho_L(x)+\rho_C(x)} dx \right) \right] dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)} \quad (\text{A9})$$

I now take the derivative with respect to $\log \sigma^\varepsilon$. This affects (A9) in two ways: both directly, via the σ^ε term in the numerator, and indirectly, via the cross-area matching rate $\rho_C(x)$ in the denominator. However, as the mobility scale σ^μ becomes large, the indirect effect via $\rho_C(x)$ loses its salience, as cross-area matching $\rho_C(x)$ becomes negligible relative to local matching $\rho_L(x)$. Therefore, as σ^μ becomes large, the derivative converges to:

$$\frac{\partial}{\partial \log \sigma^\varepsilon} \frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)} \rightarrow \pi \frac{\int_{\varepsilon}^{\infty} \left[\frac{\sigma^\varepsilon}{\sigma^\mu} \int_{\varepsilon}^{\varepsilon'} \frac{1}{r+\delta+\rho_L(x)+\rho_C(x)} dx \cdot f^\mu \left(\frac{\sigma^\varepsilon}{\sigma^\mu} \int_{\varepsilon}^{\varepsilon'} \frac{1}{r+\delta+\rho_L(x)+\rho_C(x)} dx \right) \right] dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)} \quad (\text{A10})$$

Proposition 2 states that $\frac{\partial}{\partial \log \sigma^\varepsilon} \frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)}$ is decreasing in ε , for σ^μ sufficiently large. Now, (A10) shows that $\frac{\partial}{\partial \log \sigma^\varepsilon} \frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)}$ is a weighted average of $\frac{\sigma^\varepsilon}{\sigma^\mu} \int_{\varepsilon}^{\varepsilon'} \frac{1}{r+\delta+\rho_L(x)+\rho_C(x)} dx \cdot f^\mu \left(\frac{\sigma^\varepsilon}{\sigma^\mu} \int_{\varepsilon}^{\varepsilon'} \frac{1}{r+\delta+\rho_L(x)+\rho_C(x)} dx \right)$ terms, over a truncated $F^\varepsilon(\varepsilon')$ distribution. Therefore, it is sufficient to show that $\frac{\sigma^\varepsilon}{\sigma^\mu} \int_{\varepsilon}^{\varepsilon'} \frac{1}{r+\delta+\rho_L(x)+\rho_C(x)} dx \cdot f^\mu \left(\frac{\sigma^\varepsilon}{\sigma^\mu} \int_{\varepsilon}^{\varepsilon'} \frac{1}{r+\delta+\rho_L(x)+\rho_C(x)} dx \right)$ is decreasing in ε , for every ε' , and for σ^μ sufficiently large. Equivalently, it is sufficient to show that $\mu f^\mu(\mu)$ is increasing in μ , for a positive μ which is sufficiently small.

This condition is ensured by the log concavity of f^μ , which I assume in the model. To see why, notice that an increasing $\mu f^\mu(\mu)$ is equivalent to $\frac{f^{\mu\prime}(\mu)}{f^\mu(\mu)} > -\mu$. But log concavity implies that $\frac{f^{\mu\prime}(\mu)}{f^\mu(\mu)}$ must exceed $-\infty$ as μ goes to zero (from above). Therefore, for σ^μ sufficiently large, Proposition 2 must hold.

H.4 Expression for annuitized cost $m(\mu|\varepsilon)$ in Section 4.1

In this section, I derive an expression for the annuitized cost $m(\mu|\varepsilon)$, as defined in (7). As an intermediate step, I will first solve for $\tilde{\varepsilon}(\mu|\varepsilon)$, which is the minimum match quality required for a worker (with initial match quality ε) to accept a cross-area offer with cost draw μ :

$$\tilde{\varepsilon}(\mu|\varepsilon) \equiv \varepsilon + \frac{m(\mu|\varepsilon)}{\sigma^\varepsilon} \quad (\text{A11})$$

Writing in terms of $\tilde{\varepsilon}(\mu|\varepsilon)$, worker value in (2) can be expressed as:

$$\begin{aligned} rV(\varepsilon) &= \log w(\varepsilon) + \delta[V(b) - V(\varepsilon)] + \lambda \int_\varepsilon^\infty [V(\varepsilon') - V(\varepsilon)] f^\varepsilon(\varepsilon') d\varepsilon' \quad (\text{A12}) \\ &\quad + \pi\lambda \int_0^\infty \left[\int_{\tilde{\varepsilon}(\mu|\varepsilon)}^\infty (V(\varepsilon') - V(\varepsilon) - \sigma^\mu \mu) f^\varepsilon(\varepsilon') d\varepsilon' \right] dF^\mu(\mu) \end{aligned}$$

Its derivative is:

$$V'(\varepsilon) = \frac{\sigma^\varepsilon}{r + \delta + \rho(\varepsilon)} \quad (\text{A13})$$

where $\rho(\varepsilon) \equiv \rho_L(\varepsilon) + \rho_C(\varepsilon)$ is the overall matching rate. Using (A13) and integration by parts, the value gains accruing to local and cross-area matching can be written as:

$$\int_\varepsilon^\infty [V(\varepsilon') - V(\varepsilon)] f^\varepsilon(\varepsilon') d\varepsilon' = \sigma^\varepsilon \int_\varepsilon^\infty \frac{1 - F^\varepsilon(\varepsilon')}{r + \delta + \rho(\varepsilon')} d\varepsilon' \quad (\text{A14})$$

$$\int_{\tilde{\varepsilon}(\mu|\varepsilon)}^\infty [V(\varepsilon') - V(\varepsilon) - \sigma^\mu \mu] f^\varepsilon(\varepsilon') d\varepsilon' = \sigma^\varepsilon \int_{\tilde{\varepsilon}(\mu|\varepsilon)}^\infty \frac{1 - F^\varepsilon(\varepsilon')}{r + \delta + \rho(\varepsilon')} d\varepsilon' \quad (\text{A15})$$

Substituting (A14) and (A15) into (A12), worker value can be simplified to:

$$\begin{aligned} V(\varepsilon) &= \frac{1}{r + \delta} \log w(\varepsilon) + \frac{\delta}{r + \delta} V(b) + \frac{\sigma^\varepsilon}{r + \delta} \int_\varepsilon^\infty \frac{\rho_L(\varepsilon')}{r + \delta + \rho(\varepsilon')} d\varepsilon' \quad (\text{A16}) \\ &\quad + \frac{\pi\sigma^\varepsilon}{r + \delta} \int_0^\infty \left[\int_{\tilde{\varepsilon}(\mu|\varepsilon)}^\infty \frac{\rho_L(\varepsilon')}{r + \delta + \rho(\varepsilon')} d\varepsilon' \right] dF^\mu(\mu) \end{aligned}$$

To solve for $\tilde{\varepsilon}(\mu|\varepsilon)$, I now return to the indifference condition in (7):

$$V(\tilde{\varepsilon}(\mu|\varepsilon)) = V(\varepsilon) + \sigma^\mu \mu \quad (\text{A17})$$

Replacing the value functions with (A16), I have:

$$\tilde{\varepsilon}(\mu|\varepsilon) = \varepsilon + (r + \delta) \frac{\sigma^\mu \mu}{\sigma^\varepsilon} + \int_\varepsilon^{\tilde{\varepsilon}(\mu|\varepsilon)} \frac{\rho_L(\varepsilon')}{r + \delta + \rho(\varepsilon')} d\varepsilon' + \pi \int_0^\infty \left[\int_{\tilde{\varepsilon}(x|\varepsilon)}^{\tilde{\varepsilon}(x|\tilde{\varepsilon}(\mu|\varepsilon))} \frac{\rho_L(\varepsilon')}{r + \delta + \rho(\varepsilon')} d\varepsilon' \right] dF^\mu(x) \quad (\text{A18})$$

Finally, using (A11), I can write the annuitized cost as:

$$\begin{aligned}
m(\mu|\varepsilon) &= \sigma^\varepsilon [\tilde{\varepsilon}(\mu|\varepsilon) - \varepsilon] \\
&= (r + \delta) \sigma^\mu \mu + \sigma^\varepsilon \int_\varepsilon^{\tilde{\varepsilon}(\mu|\varepsilon)} \frac{\rho_L(\varepsilon')}{r + \delta + \rho(\varepsilon')} d\varepsilon' \\
&\quad + \pi \sigma^\varepsilon \int_0^\infty \left[\int_{\tilde{\varepsilon}(x|\varepsilon)}^{\tilde{\varepsilon}(x|\tilde{\mu}(\mu|\varepsilon))} \frac{\rho_L(\varepsilon')}{r + \delta + \rho(\varepsilon')} d\varepsilon' \right] dF^\mu(x)
\end{aligned} \tag{A19}$$

To aid intuition, consider a first order approximation of $m(\mu|\varepsilon)$ around $m = 0$, as the mobility cost scale σ^μ becomes large. This yields:

$$m(\mu|\varepsilon) \approx [r + \delta + \rho_L(\varepsilon)] \sigma^\mu \mu \tag{A20}$$

which is (8) in the main text.

H.5 Proof of Proposition 3 in Section 4.3

Proposition 3 states that (i) the expected return to a cross-area match (holding human capital X_i fixed) identifies an *upper* bound on the expectation of realized annuitized costs; and (ii) the differential between the expected cross-area and same-area match returns identifies a *lower* bound on the annuitized realized costs.

The first claim is trivial: workers will only accept cross-area offers if the associated job surplus (in flow terms) exceeds the annuitized cost. I focus here on the second claim. For simplicity, I will condition my proof on individuals with some given initial match quality ε (to derive population means, one can aggregate over the ε distribution). That is, I wish to demonstrate that:

$$\int_0^\infty \left\{ \frac{\sigma^\varepsilon \int_{\varepsilon + \frac{m}{\sigma^\varepsilon}}^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon\left(\varepsilon + \frac{m}{\sigma^\varepsilon}\right)} \right\} dZ(m|\varepsilon) - \frac{\sigma^\varepsilon \int_\varepsilon^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)} \leq \int_0^\infty m dZ(m|\varepsilon) \tag{A21}$$

where the left-hand side is the differential between expected cross-area and same-area match returns, for individuals with initial match quality ε . Note the expected cross-area match return (the first term) is integrated over the distribution of realized annuitized costs, $Z(m|\varepsilon)$. The right-hand side is the expected realized cost. Equation (A21) can usefully be written as:

$$\int_0^\infty \left\{ \frac{\sigma^\varepsilon \int_{\varepsilon + \frac{m}{\sigma^\varepsilon}}^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon\left(\varepsilon + \frac{m}{\sigma^\varepsilon}\right)} - \frac{\sigma^\varepsilon \int_\varepsilon^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)} \right\} dZ(m|\varepsilon) \leq \int_0^\infty m dZ(m|\varepsilon) \tag{A22}$$

To demonstrate this claim, it is sufficient to show that:

$$\frac{\sigma^\varepsilon \int_{\varepsilon + \frac{m}{\sigma^\varepsilon}}^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon\left(\varepsilon + \frac{m}{\sigma^\varepsilon}\right)} - \frac{\sigma^\varepsilon \int_\varepsilon^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)} \leq m \tag{A23}$$

for any given annuitized cost draw $m \geq 0$. Equivalently:

$$\frac{\int_{\varepsilon + \frac{m}{\sigma^\varepsilon}}^{\infty} (\varepsilon' - \varepsilon - \frac{m}{\sigma^\varepsilon}) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})} - \frac{\int_{\varepsilon}^{\infty} (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)} \leq 0 \quad (\text{A24})$$

Since the mobility cost m must exceed zero, it is therefore sufficient to show that:

$$\frac{d}{dx} \left[\frac{\int_x^{\infty} (\varepsilon' - x) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(x)} \right] \leq 0 \quad (\text{A25})$$

for all $x \equiv \varepsilon + \frac{m}{\sigma^\varepsilon}$. Using integration by parts, (A25) is equivalent to:

$$\frac{d}{dx} \left[\frac{\int_x^{\infty} [1 - F^\varepsilon(\varepsilon')] d\varepsilon'}{1 - F^\varepsilon(x)} \right] \leq 0 \quad (\text{A26})$$

which itself is equivalent to:

$$-\frac{[1 - F^\varepsilon(x)]}{\int_x^{\infty} \left[\frac{1 - F^\varepsilon(\varepsilon')}{f^\varepsilon(\varepsilon')} \right] f^\varepsilon(\varepsilon') d\varepsilon'} + \frac{f^\varepsilon(x)}{1 - F^\varepsilon(x)} \leq 0 \quad (\text{A27})$$

Since f^ε is log concave (and therefore has a monotonically increasing hazard rate), it must be that: $\frac{f^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon')} \geq \frac{f^\varepsilon(x)}{1 - F^\varepsilon(x)}$ for all $\varepsilon' \geq x$. Therefore, $\frac{[1 - F^\varepsilon(x)]}{\int_x^{\infty} \left[\frac{1 - F^\varepsilon(\varepsilon')}{f^\varepsilon(\varepsilon')} \right] f^\varepsilon(\varepsilon') d\varepsilon'} \geq \frac{[1 - F^\varepsilon(x)]}{\int_x^{\infty} \left[\frac{1 - F^\varepsilon(x)}{f^\varepsilon(x)} \right] f^\varepsilon(\varepsilon') d\varepsilon'} = \frac{f^\varepsilon(x)}{1 - F^\varepsilon(x)}$. Plugging this back into (A27), it is clear that the equation must be satisfied.

H.6 Proof of Proposition 4 in Section 4.3

Proposition 4 states that the expected realized annuitized cost is increasing in both σ^ε and σ^μ , for given initial match quality ε . Using (13), this expectation can be written as:

$$E_Z[m|\varepsilon] = \frac{\int_0^{\infty} m [1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})] dF^m(m|\varepsilon)}{\int_0^{\infty} [1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})] dF^m(m|\varepsilon)} \quad (\text{A28})$$

The effect of mobility cost scale σ^μ is trivial. Larger σ^μ increases the annuitized cost $m(\mu|\varepsilon)$ in (A19), for given μ and ε . Hence, this pushes the annuitized cost distribution F^m to the right, so the expected cost $E_Z[m|\varepsilon]$ increases.

Next, consider the effect of match quality returns σ^ε . This enters (A28) in two ways. First, there is again an effect via the annuitized cost function $m(\mu|\varepsilon)$ in (A19), for given μ and ε (which pushes F^m to the right and increases $E_Z[m|\varepsilon]$, all else equal). Intuitively, larger σ^ε increases the value of alternative job options; so the wage gain required to accept any given cross-area offer (at non-zero cost) is larger.

Second, σ^ε also enters $E_Z[m|\varepsilon]$ directly through equation (A28). Holding the annuitized

cost distribution F^m fixed, the elasticity of $E_Z[m|\varepsilon]$ to σ^ε can be written as:

$$\frac{\partial \log E_Z[m|\varepsilon]}{\partial \log \sigma^\varepsilon} = \frac{\int_0^\infty m [1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})] \frac{\partial \log [1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})]}{\partial \log \sigma^\varepsilon} dF^m(m|\varepsilon)}{\int_0^\infty m [1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})] dF^m(m|\varepsilon)} \quad (\text{A29})$$

$$- \frac{\int_0^\infty [1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})] \frac{\partial \log [1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})]}{\partial \log \sigma^\varepsilon} dF^m(m|\varepsilon)}{\int_0^\infty [1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})] dF^m(m|\varepsilon)}$$

By inspection, a sufficient condition for a positive effect, i.e. $\frac{\partial \log E_Z[m|\varepsilon]}{\partial \log \sigma^\varepsilon} > 0$, is that the elasticity $\frac{d \log [1 - F^\varepsilon(\varepsilon + \frac{m}{\sigma^\varepsilon})]}{d \log \sigma^\varepsilon}$ is increasing in the cost draw m . And this is ensured by my assumption that the offer distribution F^ε is log concave.

H.7 Derivation of equation (15) in Section 6

Equation (15) approximates the odds ratio of cross-area to local matching in terms of the expected same-area match surplus. This approximation relies on two assumptions on the mobility cost distribution. First, I assume the draws of $\mu \sim F^\mu$ are distributed uniformly between 0 and a maximum normalized to 1. Second, I assume there are no wage offers which can justify moving at the maximum cost draw: that is, for every initial match quality ε and for every offer $\varepsilon' \sim F^\varepsilon$, $V(\varepsilon') - V(\varepsilon) < \sigma^\mu$.

Using (4), the cross-area matching rate then collapses to:

$$\rho_C(\varepsilon) = \pi \lambda \int_\varepsilon^\infty \left[\frac{\sigma^\varepsilon}{\sigma^\mu} \int_\varepsilon^{\varepsilon'} \frac{1}{r + \delta + \rho_L(x) + \rho_C(x)} dx \right] dF^\varepsilon(\varepsilon') \quad (\text{A30})$$

Linearizing around $\varepsilon' = \varepsilon$, notice that $\int_\varepsilon^{\varepsilon'} \frac{1}{r + \delta + \rho_L(x) + \rho_C(x)} dx \approx \frac{\varepsilon' - \varepsilon}{r + \delta + \rho_L(\varepsilon) + \rho_C(\varepsilon)}$. Inserting this into (A30), I have:

$$\rho_C(\varepsilon) \approx \frac{\pi \lambda}{r + \delta + \rho_L(\varepsilon) + \rho_C(\varepsilon)} \cdot \frac{\sigma^\varepsilon}{\sigma^\mu} \int_\varepsilon^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon') \quad (\text{A31})$$

Taking the ratio of (A31) to the local matching rate (3) then gives:

$$\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)} \approx \frac{\pi \sigma^\varepsilon}{[r + \delta + \rho_L(\varepsilon) + \rho_C(\varepsilon)] \sigma^\mu} \cdot \frac{\int_\varepsilon^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)} \quad (\text{A32})$$

where $\frac{\int_\varepsilon^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)}$ is the expected increase in match quality ($\varepsilon' - \varepsilon$) in a same-area match, conditional on initial quality ε . Taking expectations of (A32) over the equilibrium match quality distribution $G(\varepsilon)$ in (A6):

$$\int_b^\infty \left[\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)} \right] g(\varepsilon) d\varepsilon = \int_b^\infty \frac{\pi \sigma^\varepsilon}{[r + \delta + \rho_L(\varepsilon) + \rho_C(\varepsilon)] \sigma^\mu} \left[\frac{\int_\varepsilon^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)} \right] g(\varepsilon) d\varepsilon \quad (\text{A33})$$

where b is the reservation quality, which bounds the G distribution below. Abusing Jensen's inequality, (A33) can be approximated as:

$$\frac{\bar{\rho}_C}{\bar{\rho}_L} \approx \frac{\pi \sigma^\varepsilon}{(r + \delta + \bar{\rho}_L + \bar{\rho}_C) \sigma^\mu} \int_b^\infty \left[\frac{\int_\varepsilon^\infty (\varepsilon' - \varepsilon) dF^\varepsilon(\varepsilon')}{1 - F^\varepsilon(\varepsilon)} \right] g(\varepsilon) d\varepsilon \quad (\text{A34})$$

where $\bar{\rho}_L \equiv \int_b^\infty \rho_L(\varepsilon) g(\varepsilon) d\varepsilon$ and $\bar{\rho}_C \equiv \int_b^\infty \rho_C(\varepsilon) g(\varepsilon) d\varepsilon$ are the aggregate local and cross-area matching rates (averaged over the distribution of workers). The integral on the right-hand side of (A34) is the expected increase in match quality ε in a same-area match, now averaged across all workers. The product of this integral and match quality returns σ^ε approximates to the expected surplus in same-area matches, i.e. $\mathbb{E}[D_{ijm} - D_{ijm-1}]$ in the main text. And therefore, after taking logs, equation (15) follows from (A34).

H.8 Proof of Proposition 5 in Section 7

Proposition 5 states that, to close the mobility gap, the required subsidy s to the low- σ^ε worker is bounded above by the expected cross-area match returns $\mathbb{E}[D_{ikm} - D_{ijm-1} | k \neq j]$ of the high- σ^ε worker. To demonstrate this claim, I focus on the extreme case where the low- σ^ε worker has zero match returns (i.e. $\sigma^\varepsilon = 0$). Clearly, the required subsidy will be smaller if the low- σ^ε worker has $\sigma^\varepsilon > 0$.

To make notation clearer, I introduce new arguments into the cross-area matching rate function - for this appendix alone. Specifically, let $\rho_C(\varepsilon; \sigma^\varepsilon; s)$ denote the cross-area matching rate of a worker with initial match quality ε , match quality return σ^ε , and mobility subsidy s ; and let $F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon) | \varepsilon; \sigma^\varepsilon; s)$ be the associated distribution of annuitized mobility costs. For a worker with no subsidy, the cross-area matching rate is:

$$\rho_C(\varepsilon; \sigma^\varepsilon; 0) = \pi \lambda \int_\varepsilon^\infty F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon) | \varepsilon; \sigma^\varepsilon; 0) f^\varepsilon(\varepsilon') d\varepsilon' \quad (\text{A35})$$

which reproduces (10) in the main text. And for a worker with a subsidy s but zero match returns:

$$\rho_C(\cdot; 0; s) = \pi \lambda F^m(s | \cdot; 0; s) \quad (\text{A36})$$

Note I have replaced the ε argument with “.” in (A36): since $\sigma^\varepsilon = 0$, the worker receives no return from their ε draw; so ρ_C is independent of initial ε . The subsidy s which closes the mobility gap between the two workers satisfies:

$$\rho_C(\cdot; 0; s) = \rho_C(\varepsilon; \sigma^\varepsilon; 0) \quad (\text{A37})$$

Using (A35) and (A36), this implies:

$$F^m(s | \cdot; 0; s) = \int_\varepsilon^\infty F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon) | \varepsilon; \sigma^\varepsilon; 0) f^\varepsilon(\varepsilon') d\varepsilon' \quad (\text{A38})$$

Note that, for any given s , $F^m(s|\varepsilon; \sigma^\varepsilon; 0) < F^m(s|.; 0; s)$.³ Applying this inequality to (A38), I have:

$$F^m(s|\varepsilon; \sigma^\varepsilon; 0) < \int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon) | \varepsilon; \sigma^\varepsilon; 0) f^\varepsilon(\varepsilon') d\varepsilon' \quad (\text{A39})$$

At this stage, to ease notation, I drop the $(\varepsilon; \sigma^\varepsilon; 0)$ conditions in the F^m function; so (A39) can be simplified to:

$$F^m(s) < \int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon' \quad (\text{A40})$$

which can be rearranged to:

$$s < F^{-m} \left(\int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon' \right) \quad (\text{A41})$$

where F^{-m} is the inverse of F^m .

Proposition 5 requires that s is less than the expected cross-area match returns of the high- σ^ε worker. Therefore, to prove the proposition, it is sufficient to show that $F^{-m}(\int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon')$ on the right-hand side of (A41) is smaller than the expected returns, i.e.:

$$F^{-m} \left(\int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon' \right) < \frac{\int_{\varepsilon}^{\infty} \sigma^\varepsilon(\varepsilon' - \varepsilon) F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon'}{\int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon'} \quad (\text{A42})$$

At this stage, it is useful to define the function $H^m(m)$, where:

$$H^m(m) \equiv mF^m(m) \quad (\text{A43})$$

Using (A43), the inequality (A42) can be rewritten as:

$$F^{-m} \left(\int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon' \right) < \frac{\int_{\varepsilon}^{\infty} H^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon'}{\int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon'} \quad (\text{A44})$$

And again, using the definition of (A43), this can be expressed as:

$$H^m \left(F^{-m} \left(\int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon' \right) \right) < \int_{\varepsilon}^{\infty} H^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon' \quad (\text{A45})$$

which yields:

$$F^{-m} \left(\int_{\varepsilon}^{\infty} F^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon' \right) < H^{-m} \left(\int_{\varepsilon}^{\infty} H^m(\sigma^\varepsilon(\varepsilon' - \varepsilon)) f^\varepsilon(\varepsilon') d\varepsilon' \right) \quad (\text{A46})$$

This equation can be interpreted in terms of certainty equivalents, where the functions H^m and F^m take the place of utility functions. Specifically, the certainty equivalent of wage returns

³This is because the annuitized costs $m(\mu|\varepsilon)$ are increasing in match returns σ^ε : see equation (A19). Intuitively, larger σ^ε increases the value of alternative job options; so the wage gain required to accept any given cross-area offer (at non-zero cost) is larger.

$\sigma^\varepsilon (\varepsilon' - \varepsilon)$, over the offer distribution F^ε , is smaller when evaluated for the function F^m than for H^m . It is therefore sufficient to show that the coefficient of absolute risk aversion (CARA) is larger for F^m than H^m , i.e.:

$$-\frac{F^{m''}(m)}{F^{m'}(m)} > -\frac{H^{m''}(m)}{H^{m'}(m)} \quad (\text{A47})$$

for any given m . Replacing H^m with (A43), I have:

$$-\frac{F^{m''}(m)}{F^{m'}(m)} > -\frac{2F^{m'}(m) + mF^{m''}(m)}{F^m(m) + mF^{m'}(m)} \quad (\text{A48})$$

which rearranges to:

$$\frac{F^m(m) F^{m''}(m)}{[F^{m'}(m)]^2} < 2 \quad (\text{A49})$$

A sufficient condition for (A49), and therefore for Proposition 5, is that $f^m(m)$ is log concave. Log concavity implies that the reversed hazard rate $\frac{F^{m'}(m)}{F^m(m)}$ is monotonically decreasing in m , and this in turn implies that $\frac{F^m(m) F^{m''}(m)}{[F^{m'}(m)]^2} < 1$.

I Additional details on theoretical extensions

I.1 Endogenous search intensity

Here, I show how long-distance search intensity π can be endogenized. Suppose that, given initial match quality ε , workers choose π to maximize their value:

$$\begin{aligned} rV(\varepsilon) = & \log w(\varepsilon) + \delta [V(b) - V(\varepsilon)] + \lambda \int_{\varepsilon}^{\infty} [V(\varepsilon') - V(\varepsilon)] dF^\varepsilon(\varepsilon') \quad (\text{A50}) \\ & + \pi \lambda \int_0^{\infty} \int_{\varepsilon}^{\infty} \max \{V(\varepsilon') - V(\varepsilon) - \sigma^\mu \mu, 0\} dF^\varepsilon(\varepsilon') dF^\mu(\mu) - \frac{c}{1+\phi} \pi^{1+\phi} \end{aligned}$$

Relative to (2), I have introduced a convex long-distance search cost, $\frac{c}{1+\phi} \pi^{1+\phi}$, where $\phi > 0$. Taking the first order condition, the optimal π is then:

$$\pi(\varepsilon) = \left[\frac{\lambda}{c} \int_0^{\infty} \int_{\varepsilon}^{\infty} \max \{V(\varepsilon') - V(\varepsilon) - \sigma^\mu \mu, 0\} dF^\varepsilon(\varepsilon') dF^\mu(\mu) \right]^{\frac{1}{\phi}} \quad (\text{A51})$$

for given initial match quality ε . Equation (A51) shows the optimal π is increasing in the return to long-distance job search. This return is decreasing in ε , since workers have fewer rungs of the jobs ladder left to climb. Similarly, the optimal π is increasing in match quality returns σ^ε , which expands the expected job surplus $V(\varepsilon') - V(\varepsilon)$: see Proposition 1.

Building on (4), the rate of cross-area matching is now:

$$\rho_C(\varepsilon) = \pi(\varepsilon) \lambda \int_{\varepsilon}^{\infty} \left[F^{\mu} \left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}} \int_{\varepsilon}^{\varepsilon'} \frac{1}{r + \delta + \rho_L(x) + \rho_C(x)} dx \right) \right] dF^{\varepsilon}(\varepsilon') \quad (\text{A52})$$

where $\pi(\varepsilon)$ is determined by (A51). The main message here is that endogenous long-distance search intensity will *amplify* the impact of σ^{ε} on $\rho_C(\varepsilon)$. Proposition 1 in the baseline model describes the first order effect: σ^{ε} increases expected job surplus, which makes more long-distance matches viable. But as this appendix shows, larger surplus may also increase $\pi(\varepsilon)$, which expands $\rho_C(\varepsilon)$ still further.

I.2 Home bias in utility

I have modeled the mobility friction as a fixed one-off cost μ , which I interpret as accounting for any physical, financial, psychological or social costs of leaving home. But the basic propositions can also be derived by characterizing the friction as a home bias in utility, as is common in the urban literature: see e.g. Moretti (2011) or Diamond (2016).

More formally, suppose each worker i is assigned a “home area” h_i , and living away from home entails an amenity penalty b . A worker living in area j receives flow utility:

$$u_j(\varepsilon, d, X_i, h_i) = \log w_j(\varepsilon, X_i) - b \cdot I[j \neq h_i] \quad (\text{A53})$$

where $I[j \neq h_i]$ is an indicator function, taking 1 if the worker is away from home. The home bias $b \geq 0$ is a random draw from a distribution F^b , which is realized on arrival of a non-home offer. This variation may be motivated by the cost of distance (which varies by offer origin) or changes in personal circumstances. For simplicity, I assume there are no one-off mobility costs: all moving frictions are generated by the home bias draws b . Suppose also that workers draw new job offers from their home area j at a finite exogenous rate λ , and from elsewhere at rate $\pi\lambda$.

This yields a similar structure to the baseline model. Home area residents accept “away matches” (outside their home area) at rate:

$$\rho_A(\varepsilon) = \pi\lambda \int_{\varepsilon}^{\infty} F^b(\sigma^{\varepsilon}(\varepsilon' - \varepsilon)) dF^{\varepsilon}(\varepsilon')$$

where the home bias distribution F^b has replaced the annuitized cost distribution F^m in equation (10). Clearly, for given initial match quality ε , the away matching rate ρ_A is increasing in match quality returns σ^{ε} . This is the analogue of Proposition 1: in the face of mobility frictions (here, a consequence of home bias), larger job surplus can help justify long-distance moves.

Still, there are two differences here, which are worth noting. First, since there is no fixed cost, workers do not need to evaluate their future prospects when deciding whether to accept an away match: it is sufficient to compare the utility flows of their initial and offered jobs.

Consequently, the discount and separation rates do not matter for mobility choices. Second, this specification can motivate return migration: workers who leave home for a productive match (i.e. a high ε) may later return to enjoy home utility. [Kennan and Walker \(2011\)](#) emphasize that return migration accounts for many long-distance moves, though I show in [Table 3](#) that it does not drive education differentials in mobility.

I.3 Idiosyncratic amenity draws

In [Appendix I.2](#), workers suffer amenity penalties on leaving their home area. This can be interpreted as an environment where workers have idiosyncratic preferences over locations, and where their “home area” yields the highest idiosyncratic amenity effect. This assumption may not be unreasonable, for many individuals. Nevertheless, it is still somewhat arbitrary, especially as many long-distance moves appear to be motivated by idiosyncratic amenity *gains*: see e.g. [Appendix Table A1](#), which shows that “family” features prominently among reported reasons for moving.

To study this phenomenon, consider a more general set-up where flow utility is given by:

$$u(\varepsilon, \alpha, X_i) = \log w(\varepsilon, X_i) + \sigma^\alpha(X_i) \alpha \quad (\text{A54})$$

where α is an idiosyncratic amenity draw, specific to a worker-location match (I leave an analysis of shared non-idiosyncratic amenities to [Appendix I.4](#)). I permit dispersion σ^α in this match effect to vary across workers with different human capital X_i . Long-distance offers are characterized by both a job match quality draw $\varepsilon \sim F^\varepsilon$ and amenity draw $\alpha \sim F^\alpha$. This set-up can be interpreted as a jobs ladder in two dimensions (akin to [Sorkin, 2018](#)), namely the job match ε and amenity match α . If α is largest in a worker’s home area, setting $b = -\sigma^\alpha \alpha$ yields equation [\(A53\)](#); but this need not be true more generally.

In this more general environment, [Proposition 1](#) does not necessarily hold. Intuitively, this is because a worker already in a very high-quality job match (a consequence of large σ^ε) is less likely to be tempted by a long-distance offer with a good amenity match. To see why, notice the probability of accepting a long-distance offer with amenity draw α' (given initial job match quality ε and amenity match α) is equal to: $1 - F^\varepsilon\left(\varepsilon - \frac{\alpha' - \alpha}{\sigma^\varepsilon}\right)$. Trivially, this acceptance probability is increasing in α' ; but the sensitivity to the amenity gain $\alpha' - \alpha$ is decreasing in σ^ε . Indeed, [Panel C of Figure 2](#) shows that high-educated workers make fewer moves for non-job reasons. And [Appendix Table A2](#) shows that this effect is mainly driven by family-motivated moves, which are arguably idiosyncratic. But in practice, this negative effect is heavily dominated by the steep education gradient in job-motivated migration: see [Panel A of Figure 2](#).

Though the patterns in [Figure 2](#) are compelling, they are of course subjective reports. But, the dominance of job-motivated migration in the mobility gap is also visible in the cross-state wage returns in [Table 5](#): these are steeply *increasing* in education, consistent with selection

on large job surplus (as opposed to selection on good amenity draws⁴). This same evidence is also difficult to reconcile with an alternative hypothesis that the mobility gap is driven by education differentials in σ^α itself: i.e. that better-educated workers face larger dispersion in idiosyncratic amenity draws. Again, if high-educated workers select into mobility because of amenity gains, they should expect *lower* cross-state wage returns.

I.4 Common amenity effects

Until now, I have focused on the implications of *idiosyncratic* amenity draws α . But what happens if there is spatial dispersion in *common* amenities (such as crime and restaurants), which are shared by all workers? And what if high-educated workers care more about these common amenities, as [Diamond \(2016\)](#) suggests? In principle, differential valuation of amenities may help account for differentials in spatial mobility.⁵ Notably, [Diamond \(2016\)](#) does not actually draw this conclusion: she argues that differential amenity valuations are important for understanding the equilibrium spatial distribution of high/low-educated labor *stocks*, but she attributes differential *flows* to low-educated home bias (as in Section [I.2](#)).

Nevertheless, this hypothesis can be assessed within the framework of the paper. To this end, I replace the idiosyncratic amenity match effect α in [\(A54\)](#) with a common amenity effect a_j , specific to location j (note the distinction between the Greek α and Latin a). For a worker i living in area j , I can then write the utility flow as:

$$u_j(\varepsilon, X_i) = \log w(\varepsilon, X_i) + \sigma^a(X_i) a_j \quad (\text{A55})$$

where the wage function $w(\varepsilon, X_i)$ is defined in [\(1\)](#), and a_j is a local amenity which is common to all workers. Like [Diamond \(2016\)](#), I permit valuations of this amenity to vary across worker types: this manifests through the parameter σ^a , which depends on a worker's skill vector X_i . In the value function [\(2\)](#), this flow utility $u_j(\varepsilon, X_i)$ would then replace the $\log w(\varepsilon, X_i)$ term.

Now, consider the hypothesis that the mobility gap is driven by education variation in σ^a . This claim has two testable predictions, both of which are difficult to reconcile with the evidence in the paper. First, if high-educated workers select into mobility because of amenity gains, they should expect lower cross-state wage returns; but [Table 5](#) suggests otherwise. This is identical to the argument I apply in the case of idiosyncratic amenities: see [Appendix I.3](#). But the common amenities hypothesis also yields a second testable implication: that the mobility gap should be driven by large net flows of high-educated workers to high-amenity locations. But [Table 2](#) suggests otherwise: net flows do not explain the mobility gap.

⁴This essentially reflects the intuition of Propositions 3 and 4: if high-educated workers select into mobility because of low mobility costs, they will expect lower cross-state wage returns. Ultimately, both mobility costs and amenities are non-wage components of utility, so they have equivalent implications for cross-state wage returns.

⁵This story is closely related to the hypothesis that the mobility gap is driven by differential local returns to human capital (as in [Lkhagvasuren, 2014](#)): in each case, the mobility gap is a consequence of greater dispersion in common (i.e. non-idiosyncratic) local valuations.

References

- Diamond, Rebecca.** 2016. “The Determinants and Welfare Implications of US Workers’ Diverging Location Choices by Skill: 1980-2000.” *American Economic Review*, 106(3): 479–524.
- Faberman, R. Jason, Andreas I. Mueller, Ayşegül Şahin, and Giorgio Topa.** 2022. “Job Search Behavior Among the Employed and Non-Employed.” *Econometrica*, 90(4): 1743–1779.
- Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren.** 2018. “Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset].” University of Minnesota.
- Kaplan, Greg, and Sam Schulhofer-Wohl.** 2012. “Interstate Migration has Fallen Less than you Think: Consequences of Hot Deck Imputation in the Current Population Survey.” *Demography*, 49(3): 1061–1074.
- Kaplan, Greg, and Sam Schulhofer-Wohl.** 2017. “Understanding the Long-Run Decline in Interstate Migration.” *International Economic Review*, 58(1): 57–94.
- Kennan, John.** 2015. “Spatial Variation in Higher Education Financing and the Supply of College Graduates.” NBER Working Paper No. 21065.
- Kennan, John, and James R. Walker.** 2011. “The Effect of Expected Income on Individual Migration Decisions.” *Econometrica*, 79(1): 211–251.
- Lkhagvasuren, Damba.** 2014. “Education, Mobility and the College Wage Premium.” *European Economic Review*, 67: 159–173.
- Molloy, Raven, Christopher L. Smith, and Abigail Wozniak.** 2011. “Internal Migration in the United States.” *Journal of Economic Perspectives*, 25(3): 173–96.
- Moretti, Enrico.** 2011. “Local Labor Markets.” *Handbook of Labor Economics*, , ed. David Card and Orley Ashenfelter Vol. 4B, 1237–1313. New York: Elsevier.
- Sorkin, Isaac.** 2018. “Ranking Firms Using Revealed Preference.” *Quarterly Journal of Economics*, 133(3): 1331–1393.